Formelsammlung Statistik III Grundlagen der Statistik I: Wahrscheinlichkeitstheorie und Inferenz I

Prof. Dr. Torsten Hothorn
Institut für Statistik
Ludwig–Maximilians–Universität München
LaTeX-Satz von Dipl.-Stat. Esther Herberich

20. März 2012

Kapitel 1

Ereignisse als Mengen

1.1 Mengenoperationen

Bezeichnungen:

- $\Omega \neq \emptyset$ heißt Basismenge oder Ergebnisraum.
- Eine Menge $A \subseteq \Omega$ heißt Ereignis.
- $\omega_i \in \Omega = \{\omega_1, \omega_2, \ldots\}$ heißt Elementarereignis oder Ergebnis.
- $\mathcal{P}(\Omega) = \{A | A \subset \Omega\}$, also die Menge aller Teilmengen der Basismenge Ω, heißt Potenzmenge.
- $\mathcal{F} \subseteq \mathcal{P}(\Omega)$, also eine Menge von Teilmengen von Ω, heißt Mengensystem.

Zur Schreibweise: $\mathcal{F}, \mathcal{A}, \mathcal{E}$ Mengensysteme und A, B Mengen.

Definition 1.1 (Mengenoperationen)
Seien $A, B, A_i \subset \Omega, i \in I.$

Gleichheit:	$A = B \quad :\Leftrightarrow \quad \forall \omega \in \Omega : \quad \omega \in A \Leftrightarrow \omega \in B$
Teilmenge:	$A \subseteq B \quad :\Leftrightarrow \quad \forall \omega \in \Omega : \quad \omega \in A \Rightarrow \omega \in B$
Schnitt:	$A \cap B := \{\omega \in \Omega | (\omega \in A) \wedge (\omega \in B)\}$
	$\bigcap_{i \in I} A_i := \{\omega \in \Omega | \forall i \in I : \omega \in A_i\}$
Vereinigung:	$A \cup B := \{\omega \in \Omega | (\omega \in A) \vee (\omega \in B)\}$
	$\bigcup_{i \in I} A_i := \{\omega \in \Omega | \exists i \in I : \omega \in A_i\}$
	$\dot{\bigcup}_{i \in I} A_i := \{\omega \in \Omega | \exists i \in I : \omega \in A_i, A_i \cap A_j = \emptyset \text{ für } i \neq j\}$
Differenz:	$A \setminus B := \{\omega \in \Omega | (\omega \in A) \wedge (\omega \notin B)\}$
Komplement:	$\bar{A} := \{\omega \in \Omega | \omega \notin A\}$
Symmetrische Differenz:	$A \triangle B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$
Mächtigkeit:	$|A| :=$ Anzahl Elemente von A
Kardinalität:	$|\mathbb{N}| = \aleph_0$
	$|\mathcal{P}(\Omega)| = 2^{|\Omega|}$
Kartesisches Produkt:	$A \times B := \{(a, b) | a \in A \wedge b \in B\}$
	$\prod_{i \in I} A_i := \{(a_1, a_2, \ldots, a_{|I|}) | a_i \in A_i \forall i \in I\}$
	$A^k = \prod_{i=1,\ldots,k} A = \{(a_1, \ldots, a_k) | a_i \in A, i = 1, \ldots, k\}$

1.2 Folgen von Mengen

$(A_n)_{n\in\mathbb{N}}$

$A_n \uparrow A :\Leftrightarrow A_1 \subset A_2 \subset A_3 \subset \ldots$ und $\bigcup\limits_{i=1}^{\infty} A_i = A$

$A_n \downarrow A :\Leftrightarrow A_1 \supset A_2 \supset A_3 \supset \ldots$ und $\bigcap\limits_{i=1}^{\infty} A_i = A$

$\liminf A_n := \bigcup\limits_{n=1}^{\infty} \bigcap\limits_{m\geq n} A_n$ enthält Elemente, die in fast allen A_n enthalten sind.

$\limsup A_n := \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{m\geq n} A_n$ enthält Elemente, in unendlich vielen A_n enthalten sind.

$A = \lim_{n\to\infty} A_n \iff \liminf A_n = \limsup A_n$

Bemerkung: "fast alle" ... bis auf endlich viele

1.3 Interpretationen

$A \subset \Omega$	...	"A tritt ein", "erscheint", "wird realisiert"
$\bar{A} \subset \Omega$	...	"A tritt nicht ein"
$A_1 \cup A_2$	...	"A_1 oder A_2 treten ein"
$\bigcup\limits_{i\in I} A_i$	...	"mindestens eines der A_i tritt ein"
$A_1 \cap A_2$	...	"A_1 und A_2 treten ein"
$\bigcap\limits_{i\in I} A_i$	...	"alle A_i treten ein"
$A_1 \cup A_2 = \emptyset$	...	"A_1 und A_2 treten nicht gleichzeitig ein", "sind unvereinbar"
$A_1 \triangle A_2$	...	"Entweder A_1 oder A_2 tritt ein"
$A_1 = A_2$	...	"A_1 und A_2 beschreiben das gleiche Ereignis"
Ω	...	"Das sichere Ereignis"
$\bar{\Omega}$	...	"Das unmögliche Ereignis"

1.4 Gesetzmäßigkeiten

$A, B, C \subset \Omega$.

Reflexivität:	$A \subseteq A$
Asymmetrie:	$A \subseteq B$ und $B \subseteq A \Rightarrow A = B$
Transivität:	$A \subseteq B$ und $B \subseteq C \Rightarrow A \subseteq C$
	$\Rightarrow \mathcal{P}(\Omega)$ ist bezüglich $\subseteq$ partiell geordnet.
Kommutativgesetz:	$A \cup B = B \cup A$
	$A \cap B = B \cap A$
Assoziativgesetz:	$(A \cup B) \cup C = A \cup (B \cup C)$
	$(A \cap B) \cap C = A \cap (B \cap C)$
De Morgansche Regeln:	$\overline{A \cup B} = \bar{A} \cap \bar{B}$
	$\overline{A \cap B} = \bar{A} \cup \bar{B}$
Mächtigkeiten:	
Gleichmächtigkeit:	$\lvert A\rvert = \lvert B\rvert :\iff \exists f : A \to B$ bijektiv
Addition von Mächtigkeiten:	$A \cap B = \emptyset \iff \lvert A\rvert + \lvert B\rvert = \lvert A \cup B\rvert$

Satz 1.2 (Multiplikation von Mächtigkeiten)

$$|A \times B| = |A| \cdot |B|$$

1.5 Kombinatorik

Betrachte Indexmengen, also Teilmengen von $\{1,\ldots,n\}, n,k \in \mathbb{N}$:

Definition 1.3 (Variationen (Beachtung der Reihenfolge))

- mit Wiederholung

$$V^{\mathrm{W}}_{k,n} = \{x = (x_1,\ldots,x_k) | x \in \{1,\ldots,n\}^k\}$$

- ohne Wiederholung

$$V^{\mathrm{oW}}_{k,n} = \{x = (x_1,\ldots,x_k) | x \in \{1,\ldots,n\}^k; x_i \neq x_j \forall i \neq j\}$$

Definition 1.4 (Kombinationen (ohne Beachtung der Reihenfolge))

- mit Wiederholung

$$K^{\mathrm{W}}_{k,n} = \{x = (x_1,\ldots,x_k) | x \in \{1,\ldots,n\}^k; 1 \leq x_1 \leq \ldots \leq x_k \leq n\}$$

- ohne Wiederholung

$$K^{\mathrm{oW}}_{k,n} = \{x = (x_1,\ldots,x_k) | x \in \{1,\ldots,n\}^k; 1 \leq x_1 < \ldots < x_k \leq n\}$$

Satz 1.5

a) $|V^{\mathrm{W}}_{k,n}| = n^k$

b) $|V^{\mathrm{oW}}_{k,n}| = \frac{n!}{(n-k)!} \quad (k \leq n)$

c) $|K^{\mathrm{W}}_{k,n}| = \binom{n+k-1}{k}$

d) $|K^{\mathrm{oW}}_{k,n}| = \binom{n}{k} \quad (k \leq n)$

Satz 1.6

a) Eine n-elementige Menge $A = \{a_1,\ldots,a_n\}$ besitzt genau $\binom{n}{k}$ k-elementige Teilmengen ohne Wiederholung und ohne Beachtung der Reihenfolge.

b) Seien $j_1,\ldots,j_k \in \mathbb{N}_0$; $\sum\limits_{i=1}^{k} j_i = n$.
Es gibt genau $\binom{n}{j_1,j_2,\ldots,j_k} = \frac{n!}{j_1!j_2!\cdots j_k!}$ Möglichkeiten, $\{1,\ldots,n\}$ in eine Folge von Mengen $M_1,\ldots,M_k$; $M_i \subset \{1,\ldots,n\}$, $|M_i| = j_i$ ohne Wiederholung und ohne Beachtung der Reihenfolge aufzuteilen.

Kapitel 2

Grundlagen der Maßtheorie

2.1 Mengensysteme

Definition 2.1 (π-System)
Ein Mengensystem $\mathcal{F} \subset \mathcal{P}(\Omega)$ heißt durchschnittsstabil, wenn gilt:

$$A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}.$$

$\mathcal{F}$ heißt dann auch π-System.

Definition 2.2 (σ-Algebra)
Ein Mengensystem $\mathcal{F} \subset \mathcal{P}(\Omega)$ heißt σ-Algebra über Ω, falls gilt:

(S1) $\Omega \in \mathcal{F}$ (Basismenge)

(S2) $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$ (Komplement)

(S3) $A_i \in \mathcal{F}, i \in \mathbb{N} \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$ (abzählbare Vereinigung)

Definition 2.3 (erzeugte σ-Algebra)
Ist $A \subset \Omega$, so heißt die Menge

$$\sigma(A) = \{\emptyset, A, \bar{A}, \Omega\}$$

die von A erzeugte σ-Algebra und ist die kleinste σ-Algebra, die A enthält.

Satz 2.4 (Schnitt von σ-Algebren)
Sei $I \neq \emptyset$ eine beliebige Indexmenge und $\mathcal{F}_i$ eine σ-Algebra über Ω für alle $i \in I$. Dann ist auch

$$\mathcal{F} := \bigcap_{i \in I} \mathcal{F}_i$$

eine σ-Algebra über Ω.

Definition 2.5 (Erzeuger)
Sei $\mathcal{E} \subset \mathcal{P}(\Omega)$ ein Mengensystem und Σ die Menge aller σ-Algebren über Ω, die $\mathcal{E}$ enthalten. Dann wird die σ-Algebra

$$\sigma(\mathcal{E}) := \bigcap_{\mathcal{F} \in \Sigma} \mathcal{F}$$

als die von $\mathcal{E}$ erzeugte σ-Algebra $\sigma(\mathcal{E})$ bezeichnet. Gilt umgekehrt für eine σ-Algebra $\mathcal{A}$

$$\sigma(\mathcal{E}) = \mathcal{A}$$

so heißt $\mathcal{E}$ Erzeuger von $\mathcal{A}$.

Sei jetzt $\Omega = \mathbb{R}^n$ und $]a,b[:= \{x \in \mathbb{R} | a < x < b\}$ ein offenes, beschränktes Intervall.

Definition 2.6 (Borelsche σ-Algebra, Borelsche Mengen)
Sei $\Omega = \mathbb{R}$ und

$$\mathcal{O} = \{]a,b[| a,b \in \mathbb{R}, a < b\}$$

das Mengensystem der offenen Intervalle von $\mathbb{R}$. Dann heißt

$$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{O})$$

die Borelsche σ-Algebra über $\mathbb{R}$; ihre Elemente $A \in \mathcal{B}(\mathbb{R})$ heißen Borelsche Mengen. Für $\Omega = \mathbb{R}^n$ setzen wir

$$\mathcal{O}^n = \{U \subset \mathbb{R}^n | U \text{ offen}\} \text{ und } \mathcal{B}^n := \mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{O}^n)$$

Satz 2.7 (Eigenschaften von $\mathcal{B}$)

i) $\emptyset \in \mathcal{B}, \mathbb{R} \in \mathcal{B}$

ii) $\forall a,b \in \mathbb{R}, a < b$:
$[a,b] \in \mathcal{B}, [a,b[\in \mathcal{B},]a,b] \in \mathcal{B}$

iii) $\{c\} \in \mathcal{B} \quad \forall c \in \mathbb{R}$

iv) $\mathbb{N} \in \mathcal{B}, \mathbb{Q} \in \mathcal{B}, \bar{\mathbb{Q}} \in \mathcal{B}$

Definition 2.8 (λ-System)
Ein Mengensystem $\mathcal{D} \subset \mathcal{P}(\Omega)$ heißt λ-System über Ω, falls

(L1) $\Omega \in \mathcal{D}$

(L2) $A, B \in \mathcal{D}$ mit $A \subset B \Rightarrow B \setminus A \in \mathcal{D}$

(L3) $A_i \uparrow A \Rightarrow A \in \mathcal{D}, \quad A = \bigcup_{i=1}^{\infty} A_i$

Lemma 2.9
Ist $\mathcal{D}$ ein durchschnittsstabiles λ-System, so ist $\mathcal{D}$ eine σ-Algebra.

Satz 2.10 (λ-π-System)
Sei $\mathcal{E}$ ein π-System und $\mathcal{D}$ ein λ-System mit $\mathcal{E} \subset \mathcal{D}$. Dann gilt

$$\sigma(\mathcal{E}) \subset \mathcal{D}$$

Sei $f : \Omega_1 \to \Omega_2$ eine Abbildung, so ist das Bild einer Menge $A \subset \Omega_1$

$$f(A) := \{f(\omega) \in \Omega_2 | \omega \in A\}$$

und das Urbild einer Menge $B \subset \Omega_2$

$$f^{-1}(B) := \{\omega \in \Omega_1 | f(\omega) \in B\}.$$

Es gilt für beliebige $B, B_i \in \Omega_2, i \in I$ ("Operationstreue")

$$\begin{aligned}
f^{-1}\left(\bigcup_{i\in I} B_i\right) &= \left\{\omega \in \Omega_1 | f(\omega) \in \bigcup_{i\in I} B_i\right\} \\
&= \bigcup_{i\in I} \{\omega \in \Omega_1 | f(\omega) \in B_i\} \\
&= \bigcup_{i\in I} f^{-1}(B_i) \\
f^{-1}\left(\bigcap_{i\in I} B_i\right) &= \bigcap_{i\in I} f^{-1}(B_i) \\
f^{-1}(\bar{B}) &= \{\omega \in \Omega_1 | f(\omega) \in \bar{B}\} \\
&= \Omega_1 \setminus \{\omega \in \Omega_1 | f(\omega) \in B\} \\
&= \Omega_1 \setminus f^{-1}(B) = \overline{f^{-1}(B)}
\end{aligned}$$

Definition 2.11
Sei $\mathcal{F} \subset \mathcal{P}(\Omega)$ ein Mengensystem:

$$f^{-1}(\mathcal{F}) = \{f^{-1}(B) | B \in \mathcal{F}\}.$$

Satz 2.12
Sei $\mathcal{F}_i$ eine σ-Algebra über $\Omega_i, i = 1, 2$. Ist $f : \Omega_1 \to \Omega_2$ eine Abbildung, so ist $f^{-1}(\mathcal{F}_2)$ eine σ-Algebra über Ω_1 und $\{B \subset \Omega_2 | f^{-1}(B) \in \mathcal{F}_1\}$ eine σ-Algebra über Ω_2.

2.2 Meßbare Abbildungen

Definition 2.13 (meßbar)
Sei $\mathcal{F}$ eine σ-Algebra über Ω und $A \in \mathcal{F}$. Dann heißt A <u>meßbar</u> bezüglich $\mathcal{F}$.

Definition 2.14 (Meßraum)
Sei $\mathcal{F}$ eine σ-Algebra über Ω. Das Tupel $(\Omega, \mathcal{F})$ heißt <u>Meßraum</u>.

Definition 2.15 (meßbare Abbildung)
Seien $(\Omega_1, \mathcal{F}_1)$ und $(\Omega_2, \mathcal{F}_2)$ zwei Meßräume. Eine Abbildung $f : \Omega_1 \to \Omega_2$ heißt <u>$\mathcal{F}_1$-$\mathcal{F}_2$-meßbar</u>, falls

$$f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1$$

. Von einer meßbaren Abbildung f spricht man, falls die involvierten σ-Algebren eindeutig bekannt sind.

Satz 2.16
Seien $(\Omega_i, \mathcal{F}_i), i = 1, 2$ zwei Meßräume, wobei $\mathcal{F}_2 = \sigma(\mathcal{E})$ von einem Mengensystem $\mathcal{E}$ erzeugt ist. Die Abbildung $f : \Omega_1 \to \Omega_2$ ist genau dann $\mathcal{F}_1$-$\mathcal{F}_2$-meßbar, wenn

$$f^{-1}(\mathcal{E}) \subset \mathcal{F}_1$$

.

Satz 2.17 (Meßbarkeit der Komposition)
Sind $(\Omega_i, \mathcal{F}_i), i = 1, 2, 3$ Meßräume und $f : \Omega_1 \to \Omega_2$ und $g : \Omega_2 \to \Omega_3$ meßbar. Dann ist auch $g \circ f : \Omega_1 \to \Omega_3$ meßbar.

Satz 2.18
Es sei $(\Omega, \mathcal{F})$ ein Meßraum, $f, g : \Omega \to \mathbb{R}$ meßbare Funktionen und $\alpha, \beta \in \mathbb{R}$. Dann sind die Funktionen

a) $\alpha f + \beta g$

b) $f \cdot g$

c) f/g falls $g(\omega) \neq 0 \quad \forall \omega \in \Omega$

ebenfalls meßbar.

Satz 2.19
Sei $(\Omega, \mathcal{F})$ ein Meßraum und $(f_n)_{n\in\mathbb{N}}$ eine Folge meßbarer numerischer Funktionen

$$f_n : \Omega \to \bar{\mathbb{R}}, n \in \mathbb{N}.$$

Dann sind $\sup\limits_n f_n$, $\inf\limits_n f_n$, $\limsup\limits_n f_n$ und $\liminf\limits_n f_n$ meßbar.

Satz 2.20
Ist $(\Omega, \mathcal{F})$ ein Meßraum und $f : \Omega \to \bar{\mathbb{R}}$ eine meßbare numerische Funktion. Dann sind auch

a) $f^+ := \max(f, 0)$ Positivteil

b) $f^- := \max(-f, 0)$ Negativteil

c) $|f| := f^+ + f^-$ Betrag

meßbar. (Bemerkung: $f = f^+ - f^-$)

2.3 Maße und deren Eigenschaften

Definition 2.21 ((σ-endliches, endliches, normiertes) Maß)
Sei $(\Omega, \mathcal{F})$ ein Meßraum. Eine Funktion $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ heißt Maß auf $\mathcal{F}$, falls die folgenden Bedingungen erfüllt sind

(M1) $\mu(\emptyset) = 0$

(M2) $\mu(A) \geq 0 \quad \forall A \in \mathcal{F}$

(M3) für jede Folge disjunkter Mengen $(A_n)_{n\in\mathbb{N}} \in \mathcal{F}$ gilt

$$\mu\left(\dot{\bigcup_{n=1}^{\infty}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \quad (\sigma\text{-Additivität})$$

Gibt es eine Folge (A_n) von Mengen aus $\mathcal{F}$ mit $\bigcup\limits_{n=1}^{\infty} A_n = \Omega$ und $\mu(A_n) < \infty \quad \forall n \in \mathbb{N}$, so heißt μ σ-endlich. Ist $\mu(\Omega) < \infty$, so heißt μ endlich. Ist $\mu(\Omega) = 1$, so heißt μ normiertes Maß.

Definition 2.22 (Maßraum)
Ist $(\Omega, \mathcal{F})$ ein Meßraum und $\mu : \mathcal{F} \to \bar{\mathbb{R}}$ ein Maß, so heißt das Tripel $(\Omega, \mathcal{F}, \mu)$ Maßraum.

Satz 2.23 (Eigenschaften des Maßes)
Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $A, B, A_n \in \mathcal{F}, n \in \mathbb{N}$. Dann gilt

i) endliche Additivität: $A \cap B = \emptyset \Rightarrow \mu(A \cup B) = \mu(A) + \mu(B)$

ii) Subadditivität: $A \subset B$ und $\mu(A) < \infty \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A)$

iii) Monotonie: $A \subset B \Rightarrow \mu(A) \leq \mu(B)$

iv) Sub-σ-Additivität:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

Satz 2.24 (Stetigkeit des Maßes μ)
Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $A, A_n \in \mathcal{F}, n \in \mathbb{N}$. Dann gilt

i) Stetigkeit von unten: $A_n \uparrow A \Rightarrow \mu(A_n) \uparrow \mu(A)$

ii) Stetigkeit von oben: $A_n \downarrow A$ und $\mu(A_1) < \infty \Rightarrow \mu(A_n) \downarrow \mu(A)$

Satz 2.25 (Maßeindeutigkeitssatz)
Es seien μ und ν zwei Maße auf einem Meßraum $(\Omega, \mathcal{F})$ und $\mathcal{E}$ ein durchschnittsstabiler Erzeuger von $\mathcal{F}$ mit folgenden Eigenschaften:

i) $\mu(E) = \nu(E) \quad \forall E \in \mathcal{E}$

ii) Es gibt eine Folge $(E_n), n \in \mathbb{N}$, disjunkter Mengen aus $\mathcal{E}$ mit $\mu(E_n) = \nu(E_n) < \infty$ und

$$\dot{\bigcup_{n=1}^{\infty}} E_n = \Omega.$$

Dann sind beide Maße identisch: $\mu = \nu$.

Definition 2.26 (Nullmenge)
Ist $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $A \in \mathcal{F}$ mit $\mu(A) = 0$, so heißt A (μ-)Nullmenge.

Definition 2.27 (vollständiges Maß, Vervollständigung)
Es sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Ein Maß μ heißt vollständig, wenn gilt:

$$A \in \mathcal{F}, \mu(A) = 0 \text{ und } B \subset A$$

$$\Rightarrow B \in \mathcal{F}.$$

Die σ-Algebra $\mathcal{F}_0 := \{A \cup N | A \in \mathcal{F}, N \text{ Teilmenge einer } \mu\text{-Nullmenge}\}$ heißt Vervollständigung von $\mathcal{F}$.

2.4 Das Lebesgue-Maß

Satz 2.28 (Existenz und Eindeutigkeit des Lebesgue-Maßes)
In jeder Dimension $n \in \mathbb{N}$ gibt es genau ein Maß

$$\lambda^n : \mathcal{B}^n \to [0, \infty[,$$

sodaß für jedes n-dimensionale Intervall $]a, b] =]a_1, b_1] \times \ldots \times]a_n, b_n] \subset \mathbb{R}^n$ gilt:

$$\lambda^n(]a, b]) = \prod_{i=1}^{n} (b_i - a_i)$$

λ^n heißt (n-dimensionales) Lebesgue-Maß. Weiterhin gibt es zu jeder rechtsseitig stetigen, monoton wachsenden Funktion $F : \mathbb{R} \to \mathbb{R}$ genau ein Maß

$$\lambda_F : \mathcal{B} \to \bar{\mathbb{R}},$$

sodaß für alle $]a, b] \subset \mathbb{R}$ gilt:

$$\lambda_F(]a, b]) = F(b) - F(a).$$

λ_F heißt Lebesgue-Stieltjes-Maß von F.

Satz 2.29 (Translationsinvarianz des Lebesgue-Maßes)

$$\lambda^n(A) = \lambda^n(A+v) \quad \forall v \in \mathbb{R}^n \quad A =]a,b]$$

Definition 2.30 (Bildmaß)
Ist $(\Omega_1, \mathcal{F}_1, \mu)$ ein Maßraum und $(\Omega_2, \mathcal{F}_2)$ Meßraum und $f : \Omega_1 \to \Omega_2$ meßbar, so heißt

$$\mu_f : \mathcal{F}_2 \to [0, \infty[$$

$$B \mapsto \mu_f(B) = \mu(f^{-1}(B))$$

das Bildmaß μ_f von μ unter f und $(\Omega_2, \mathcal{F}_2, \mu_f)$ Maßraum.

Kapitel 3

Das Lebesgue-Integral

3.1 Integral für Treppenfunktionen

Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $f : \Omega \to \mathbb{R}$ eine $\mathcal{F}$-$\mathcal{B}$-meßbare Funktion. Sei weiterhin $M := \{f | f : \Omega \to \bar{\mathbb{R}}, f \text{ meßbar}\}$ die Menge der meßbaren numerischen Funktionen und $M^+ := \{f | f \in M, f \geq 0\}$ die nicht-negativen Funktionen aus M.

Definition 3.1 (Treppenfunktion)
Ist $(\Omega, \mathcal{F}, \mu)$ ein Maßraum und $f : \Omega \to \mathbb{R}$ eine $\mathcal{F}$-$\mathcal{B}$-meßbare Funktion mit endlichem Bild $f(\Omega) = \{y_1, \ldots, y_n\}$, so läßt sich f für ein $n \in \mathbb{N}$ und $A_i \in \mathcal{F}, i = 1, \ldots, n$, darstellen als

$$f = y_1 I_{A_1} + \ldots + y_n I_{A_n} = \sum_{i=1}^{n} y_i I_{A_i}$$

und heißt Treppenfunktion. Desweiteren sei $T := \{f | f : \Omega \to \mathbb{R},\ f \text{ Treppenfunktion}\}$ und $T^+ := \{f | f \in T, f \geq 0\}$ die Menge der (nicht-negativen) Treppenfunktionen.

Definition 3.2 (Lebesgue-Integral für Treppenfunktionen)
Das Integral für $f \in T^+$

$$\int f \, d\mu := y_1 \mu(A_1) + \ldots + y_n \mu(A_n)$$

heißt Lebesgue-Integral von f nach μ.

Satz 3.3 (Eigenschaften des Integrals) i) Linearität: Für $f, g \in T^+$ und $\alpha, \beta \geq 0$ gilt

$$\int (\alpha f + \beta g) \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu.$$

ii) Monotonie: Sind $f, g \in T^+$ und $f \leq g$, so folgt

$$\int f \, d\mu \leq \int g \, d\mu.$$

3.2 Integral nicht-negativer Funktionen

Lemma 3.4
Ist $f \in M^+$ eine nicht-negative Funktion, so gibt es eine Folge $(f_n)_{n \in \mathbb{N}} \in T^+$ von nicht-negativen Treppenfunktionen, so daß $f_n \uparrow f$.

Definition 3.5 (Lebesgue-Integral für nicht-negative Funktionen)
Sei $f \in M^+$ und $(f_n)_{n\in\mathbb{N}} \in T^+$ mit $f_n \uparrow f$. Das Lebesgue-Integral von f nach μ ist dann

$$\int f\,d\mu := \lim_{n\to\infty} \int f_n\,d\mu$$

Lemma 3.6
Sei $f \in M^+$ und $g \in T^+$ mit $g \le f$. Dann gilt für eine Folge $(f_n)_{n\in\mathbb{N}} \in T^+$:

$$f_n \uparrow f \Rightarrow \lim_{n\to\infty} \int f_n\,d\mu \ge \int g\,d\mu.$$

Lemma 3.7
Sind $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ zwei monoton wachsende Folgen von Funktionen aus T^+ mit

$$\lim_{n\to\infty} f_n = \lim_{n\to\infty} g_n,$$

so gilt:

$$\lim_{n\to\infty} \int f_n\,d\mu = \lim_{n\to\infty} \int g_n\,d\mu$$

Satz 3.8 (Eigenschaften des Integrals)
Für $f, g \in M^+$ und $\alpha, \beta \ge 0$ gilt:

i) Linearität:
$$\int (\alpha f + \beta g)\,d\mu = \alpha \int f\,d\mu + \beta \int g\,d\mu$$

ii) Monotonie: Ist $f \le g$ so folgt
$$\int f\,d\mu \le \int g\,d\mu$$

Satz 3.9 (Satz von der monotonen Konvergenz, Satz von Beppo Levi)
Für eine monoton wachsende Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen aus M^+ gilt:

$$\lim_{n\to\infty} \int f_n\,d\mu = \int \lim_{n\to\infty} f_n d\mu$$

Korollar 3.10 (zu Satz 3.9)
Für jede Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen aus M^+ gilt:

$$\int \sum_{i=1}^{\infty} f_i d\mu = \sum_{i=1}^{\infty} \int f_i d\mu.$$

Satz 3.11 (Lemma von Fatou)
Für jede Folge $(f_n)_{n\in\mathbb{N}} \in M^+$ gilt:

$$\liminf_{n\to\infty} \int f_n\,d\mu \ge \int \liminf_{n\to\infty} f_n\,d\mu.$$

3.3 Integrierbare Funktionen

Jetzt allgemein $f : \Omega \to \bar{\mathbb{R}}$ meßbar.

$$f = \underbrace{f^+}_{\in M^+} - \underbrace{f^-}_{\in M^+}$$

Definition 3.12 (quasi-integrierbar, Lebesgue-Integral)
Sei $f : \Omega \to \bar{\mathbb{R}}$ eine meßbare numerische Funktion. Die Funktion heißt (μ-) quasi-integrierbar, falls $\int f^+ \, d\mu < \infty$ *oder* $\int f^- \, d\mu < \infty$. Ist f (μ-) quasi-integrierbar, so ist durch

$$\int f \, d\mu := \int f^+ \, d\mu - \int f^- \, d\mu$$

das Lebesgue-Integral von f definiert. Gilt $\int f^+ \, d\mu < \infty$ und $\int f^- \, d\mu < \infty$, so heißt f (μ-)integrierbar.

Satz 3.13
Sind $f, g : \Omega \to \bar{\mathbb{R}}$ integrierbare numerische Funktionen und $\alpha, \beta \in \mathbb{R}$, so gilt

i) Linearität: $\alpha f + \beta g$ ist integrierbar und

$$\int (\alpha f + \beta g) \, d\mu = \alpha \int f \, d\mu + \beta \int g \, d\mu$$

ii) Monotonie: Ist $f \leq g$, so folgt

$$\int f \, d\mu \leq \int g \, d\mu$$

iii) Für jedes $A \in \mathcal{F}$ gilt

$$\int f \, d\mu = \int_A f \, d\mu + \int_{\bar{A}} f \, d\mu$$

Satz 3.14 (Satz von der dominierten Konvergenz)
Seien f, g sowie $(f_n), (g_n)$ meßbare numerische Funktionen auf einem Maßraum $(\Omega, \mathcal{F}, \mu)$ mit $f_n \to f, g_n \to g$ (punktweise) und $|f_n| \leq g_n \quad \forall n \in \mathbb{N}$. Sind g und $g_n \quad \forall n \in \mathbb{N}$ integrierbar und gilt

$$\lim_{n \to \infty} \int g_n \, d\mu = \int g \, d\mu,$$

so sind f und $f_n \quad \forall n \in \mathbb{N}$ integrierbar und es gilt

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu.$$

Satz 3.15
Für $f \in M^+$ gilt

$$\int f \, d\mu = 0 \Leftrightarrow \underbrace{\{\omega \in \Omega | f(\omega) > 0\}}_{=: \text{ Träger von } f} \text{ ist } \mu\text{-Nullmenge}$$

Definition 3.16 (μ-fast-überall)
Die Eigenschaft E sei für die Elemente $\omega \in \Omega$ eines Maßraumes $(\Omega, \mathcal{F}, \mu)$ sinnvoll. E gilt (μ-)fast-überall, wenn E für alle $\omega \notin N \subset \Omega$ gilt und N eine μ-Nullmenge ist.

Satz 3.17 (Riemann & Lebesgue-Integral)
Sei $f : \mathbb{R} \to \mathbb{R}$ meßbar und auf $[a, b] \subset \mathbb{R}$ Riemann-integrierbar (Ober- und Untersumme konvergieren gegen Integralwert). Dann ist f Lebesgue-integrierbar und es gilt

$$\underbrace{\int_{[a,b]} f \, d\lambda}_{\text{Lebesgue}} = \underbrace{\int_a^b f(x) \, dx}_{\text{Riemann}}$$

Definition 3.18 (Produktmaßraum)
Seien $(\Omega_1, \mathcal{F}_1, \mu)$ und $(\Omega_2, \mathcal{F}_2, \nu)$ zwei Maßräume. Dann heißt der Maßraum

$$(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \mu \otimes \nu)$$

mit Basismenge

$$\Omega_1 \times \Omega_2,$$

Produkt-σ-Algebra

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\{A \times B | A \in \mathcal{F}_1, B \in \mathcal{F}_2\})$$

(wobei $\{A \times B | A \in \mathcal{F}_1, B \in \mathcal{F}_2\}$ ein durchschnittsstabiler Erzeuger von $\mathcal{F}_1 \otimes \mathcal{F}_2$ ist)
und Produktmaß

$$\mu \otimes \nu(A \times B) = \mu(A) \cdot \nu(B) \quad \forall A \in \mathcal{F}_1, B \in \mathcal{F}_2$$

<u>Produktmaßraum</u>.

Satz 3.19 (Satz von Fubini)
Seien $(\Omega_1, \mathcal{F}_1, \mu)$ und $(\Omega_2, \mathcal{F}_2, \nu)$ zwei Maßräume mit σ-endlichen Maßen μ und ν. Ist $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ eine nicht-negative, $(\mathcal{F}_1 \otimes \mathcal{F}_2)$-$\mathcal{B}$-meßbare Funktion oder eine $(\mu \otimes \nu)$-integrierbare Funktion, so gilt

$$\begin{aligned}\int f \, d(\mu \otimes \nu) &= \int \left(\int f(\omega_1, \omega_2) \, d\mu(\omega_1) \right) d\nu(\omega_2) \\ &= \int \left(\int f(\omega_1, \omega_2) \, d\nu(\omega_2) \right) d\mu(\omega_1)\end{aligned}$$

(Vertauschung der Reihenfolge der Integration).

3.4 Ungleichungen

$f : \Omega \to \mathbb{R}$ $\mathcal{F}$-$\mathcal{B}$-meßbar auf $(\Omega, \mathcal{F}, \mu)$

$$\|f\|_p := \left(\int |f|^p d\mu \right)^{1/p} \in [0, \infty[\quad \text{"}p\text{-Norm"}$$

Satz 3.20 (Ungleichung von Hölder)
Es sei $1 < p, q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für zwei meßbare Funktionen $f, g : \Omega \to \mathbb{R}$:

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q.$$

Satz 3.21 (Ungleichung von Minkowski)
Sind $f, g : \Omega \to \mathbb{R}$ meßbar und $p \geq 1$, so gilt

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad \text{Dreiecksungleichung}$$

Definition 3.22 ($\mathcal{L}^p$-Raum)
Für $p \geq 1$ sei

$$\mathcal{L}^p := \mathcal{L}^p(\Omega, \mathcal{F}, \mu) := \{f | f : \Omega \to \mathbb{R}, f \text{ meßbar}, \|f\|_p < \infty\}$$

Satz 3.23
Ist $\mu(\Omega) < \infty$ und $q > p \geq 1$, so ist $L^q \subset L^p$ und es gibt $c \geq 0$, so daß

$$\|f\|_p \leq c\|f\|_q \quad \forall f \in L^q.$$

Definition 3.24 (Konvergenz in L^p)
Eine Folge $(f_n)_{n\in\mathbb{N}} \in L^p$ heißt in L^p konvergent, wenn es ein $f \in L^p$ gibt, so daß

$$\|f_n - f\|_p \to 0 \quad \text{für } n \to \infty$$

Kurz: $f_n \overset{L^p}{\to} f$.

Definition 3.25 (Konvergenz μ-f.ü.)
Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Eine Folge $f_n : \Omega \to \mathbb{R}$ meßbarer Funktionen heißt (μ-) fast überall konvergent, wenn es eine meßbare Funktion $f : \Omega \to \mathbb{R}$ und eine μ-Nullmenge N gibt, so daß

$$f_n(x) \to f(x) \quad \forall x \in \bar{N}.$$

Kurz: $f_n \to f$ μ-f.ü.

Korollar 3.26 (zu Satz 3.23)
Ist $q > p \geq 1,\ f, f_n \in L^q,\ n \in \mathbb{N}$, so gilt:

$$f_n \overset{L^q}{\to} f \Rightarrow f_n \overset{L^p}{\to} f$$

Satz 3.27
Seien $p \geq 1$ und $f, f_n \in L^p,\ n \in \mathbb{N}$, so daß $f_n \to f$ f. ü. Dann gilt

$$\|f_n\|_p \to \|f\|_p \Leftrightarrow f_n \overset{L^p}{\to} f.$$

3.5 Maße und Dichten

Lemma 3.28
Für jedes $f \in M^+$ ist

$$f \odot \mu : \mathcal{F} \to \bar{\mathbb{R}}$$
$$(f \odot \mu)(A) := \int_A f\, d\mu = \int f I_A\, d\mu$$

ein Maß. Sei $\nu = f \odot \mu$, so gilt für alle $g \in M^+$

$$\int g\, d\nu = \int (gf)\, d\mu.$$

Definition 3.29 (Maß mit Dichte, Dichte)
Sei $(\Omega, \mathcal{F}, \mu)$ Maßraum und $f \in M^+$. Dann heißt $f \odot \mu$ Maß mit der Dichte f bezüglich μ. Die Funktion $f : \Omega \to [0, \infty[$ heißt Dichte des Maßes $f \odot \mu$.

Definition 3.30 (absolute Stetigkeit)
Sind μ und ν zwei Maße auf $(\Omega, \mathcal{F})$, so heißt ν absolut stetig bezüglich μ, wenn $\forall A \in \mathcal{F}$ gilt:

$$\mu(A) = 0 \Rightarrow \nu(A) = 0$$

Kurz: $\nu \ll \mu$. Man sagt auch: μ dominiert ν.

Satz 3.31 (Satz von Radon-Nikodym)
Ist $(\Omega, \mathcal{F}, \mu)$ ein Maßraum mit einem σ-endlichen Maß μ und ν ein weiteres Maß auf $(\Omega, \mathcal{F})$, gilt

$$\nu \ll \mu \Leftrightarrow \exists \text{ Dichte } f \in M^+ : \nu = f \odot \mu.$$

Gibt es eine weitere Funktion $g \in M^+$ mit $\nu = g \odot \mu$ so gilt $g = f$ μ-f.ü.

Kapitel 4

Wahrscheinlichkeitsmaße

4.1 Wahrscheinlichkeit als normiertes Maß

Jetzt Ω Ergebnisraum. Potentiell interessante Ereignisse sind Teilmengen von Ω, d.h., Elemente einer sinnvollen σ-Algebra $\mathcal{F}$, des Ereignisraumes.

Definition 4.1 (Wahrscheinlichkeitsmaß, Wahrscheinlichkeit)
Sei $(\Omega, \mathcal{F})$ ein Meßraum und $\mathbb{P} : \mathcal{F} \to [0,1]$ ein Maß auf $(\Omega, \mathcal{F})$ mit $\mathbb{P}(\Omega) = 1$. Dann heißt $\mathbb{P}$ Wahrscheinlichkeitsmaß und ordnet jedem Ereignis $A \in \mathcal{F}$ die Wahrscheinlichkeit $\mathbb{P}(A)$ zu.

Definition 4.2 (Wahrscheinlichkeitsraum, Verteilung)
$(\Omega, \mathcal{F}, \mathbb{P})$ heißt Wahrscheinlichkeitsraum. $\mathbb{P}$ heißt auch Verteilung.

Definition 4.3 ($\mathbb{P}$-fast sicher)

$$\mathbb{P}\text{-fast überall} \Leftrightarrow \mathbb{P}\text{-fast sicher}$$

Satz 4.4 (Elementare Rechenregeln)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A, B, A_n \in \mathcal{F}$, $n \in \mathbb{N}$. Dann gilt

i) $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$

ii) $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

iii) Siebformel:
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} (-1)^{i+1} \sum_{1 \leq j_1 < \ldots < j_i \leq n} \mathbb{P}\left(\bigcap_{k=1}^{i} A_{j_k}\right)$$

iv) $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

v) Stetigkeit von unten: $A_n \uparrow A \Rightarrow \mathbb{P}(A_n) \uparrow \mathbb{P}(A)$

vi) Stetigkeit von oben: $A_n \downarrow A \Rightarrow \mathbb{P}(A_n) \downarrow \mathbb{P}(A)$

4.2 Diskrete Wahrscheinlichkeitsmaße

Definition 4.5 (diskreter W'keitsraum)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Gibt es eine abzählbare (endlich oder unendlich) Menge $T \in \mathcal{F}$ mit $\mathbb{P}(T) = 1$, so heißt $(\Omega, \mathcal{F}, \mathbb{P})$ diskreter Wahrscheinlichkeitsraum, $\mathbb{P}$ diskretes Wahrscheinlichkeitsmaß und T ein abzählbarer Träger von $\mathbb{P}$.

Satz 4.6
Ist $(\Omega, \mathcal{F}, \mathbb{P})$ ein diskreter Wahrscheinlichkeitsraum mit Träger T, so ist die Funktion

$$\begin{aligned} f : \Omega &\to [0,1] \\ \omega &\mapsto f(\omega) = \begin{cases} \mathbb{P}(\{\omega\}) & \text{für } \omega \in T \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

eine Dichte von $\mathbb{P}$ bezüglich des Zählmaßes μ_Z, also $\mathbb{P} = f \odot \mu_Z$. Insbesondere gilt

$$\mathbb{P}(A) = \int_A f \, d\mu_Z = \sum_{\omega \in A \cap T} f(\omega) \quad \forall A \in \mathcal{F}. \tag{4.1}$$

Umgekehrt gibt es zu jeder Funktion f der obigen Gestalt genau ein diskretes Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $(\Omega, \mathcal{F})$, so daß (4.1) gilt.

Definition 4.7 (Zähldichte)
f aus Satz 4.6 heißt Zähldichte.

4.3 Stetige Verteilungen

Jetzt: $\mathbb{P} : \mathcal{B} \to [0,1]$

Definition 4.8 (Verteilungsfunktion)
Ist $\mathbb{P} : \mathcal{B} \to [0,1]$ ein Wahrscheinlichkeitsmaß auf $\mathbb{R}$, so heißt

$$\begin{aligned} F_{\mathbb{P}} : \mathbb{R} &\to [0,1] \\ x &\mapsto F_{\mathbb{P}}(x) := \mathbb{P}(]-\infty, x]) \end{aligned}$$

die Verteilungsfunktion von $\mathbb{P}$.

Satz 4.9 (Eigenschaften der Verteilungsfunktion)
Eine Verteilungsfunktion $F_{\mathbb{P}}$ ist

i) monoton wachsend

ii) rechtsstetig

iii) $\lim\limits_{x \to -\infty} F_{\mathbb{P}}(x) = 0; \quad \lim\limits_{x \to \infty} F_{\mathbb{P}}(x) = 1$

Definition 4.10 (Verteilungsfunktion)
Jede monoton wachsende, rechtsstetige Funktion

$$F : \mathbb{R} \to [0,1] \quad \text{mit}$$

$$\lim_{x \to -\infty} F(x) = 0 \;\text{ und }\; \lim_{x \to \infty} F(x) = 1$$

heißt Verteilungsfunktion.

Satz 4.11 (Korrespondenzsatz)
Für jede Verteilungsfunktion F ist $\mu := \lambda_F(]a,b]) = F(b) - F(a)$ (siehe Satz 2.28) ein Wahrscheinlichkeitsmaß mit $F_\mu = F$. Umgekehrt ist für jedes reelle Wahrscheinlichkeitsmaß $\mathbb{P}$ die Funktion $G := F_{\mathbb{P}}$ eine Verteilungsfunktion und $\lambda_G = \mathbb{P}$.

Lemma 4.12
Sei $\mathbb{P} : \mathcal{B} \to [0,1]$ ein Wahrscheinlichkeitsmaß und $F_{\mathbb{P}}$ seine Verteilungsfunktion. Dann gilt

$$\mathbb{P}(\{x\}) = F_{\mathbb{P}}(x) - F_{\mathbb{P}}(x^-) \quad \forall x \in \mathbb{R}$$

mit $F_{\mathbb{P}}(x^-) = \lim\limits_{t \uparrow x} F(t) \quad \forall x \in \mathbb{R}$ (linksseitiger Grenzwert)

Korollar 4.13
Sei $\mathbb{P} : \mathcal{B} \to [0,1]$ ein Wahrscheinlichkeitsmaß und $F_{\mathbb{P}}$ seine Verteilungsfunktion. Dann sind äquivalent

i) $F_{\mathbb{P}}$ ist stetig

ii) $\mathbb{P}(\{x\}) = 0 \quad \forall x \in \mathbb{R}$.

4.4 Dichten

Sei $\mu = (f \odot \lambda)$ ein Maß mit der Dichte f bezüglich des Lebesgue-Maßes λ, also

$$\mu(A) = \int_A f \, d\lambda \quad \forall A \in \mathcal{B}.$$

Satz 4.14
Sei $f : \mathbb{R} \to [0, \infty[$ und $\mathbb{P} = f \odot \lambda$. Dann ist $\mathbb{P}$ ein Wahrscheinlichkeitsmaß genau dann, wenn f auf $\mathbb{R}$ integrierbar ist und $\displaystyle\int f \, d\lambda = 1$.

Kapitel 5

Zufallsvariablen

5.1 Zufallsvariable und deren Verteilung

Definition 5.1 (Zufallsvariable)
Ist $(\Omega_1, \mathcal{F}_1, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\Omega_2, \mathcal{F}_2)$ ein Meßraum, so heißt eine $\mathcal{F}_1$-$\mathcal{F}_2$ meßbare Abbildung

$$X : \Omega_1 \to \Omega_2$$

Zufallsvariable (ZV). Ist $\Omega_2 = \mathbb{R}$, so heißt X reelle Zufallsvariable, $\Omega_2 = \bar{\mathbb{R}}$ numerische Zufallsvariable, $\Omega_2 = \mathbb{R}^n$ n-dimensionale reelle Zufallsvariable.

Definition 5.2 (Verteilung einer Zufallsvariablen)
Ist $(\Omega_1, \mathcal{F}_1, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $(\Omega_2, \mathcal{F}_2)$ ein Meßraum und $X : \Omega_1 \to \Omega_2$ eine Zufallsvariable, so heißt das Bildmaß $\mathbb{P}_X$ von $\mathbb{P}$ unter X Verteilung von X.

Satz 5.3
Sei $X : \Omega \to \mathbb{R}^n$ eine n-dimensionale reelle Zufallsvariable und $f : \mathbb{R}^n \to \mathbb{R}$ eine meßbare Funktion, so daß $f \circ X : \Omega \to \mathbb{R}$ integrierbar ist. Dann gilt

$$\int f \, d\mathbb{P}_X = \int f \circ X \, d\mathbb{P}.$$

5.2 Erwartungswert und Varianz

Definition 5.4 (Erwartungswert)
Ist $X : \Omega \to \mathbb{R}$ eine quasi-integrierbare reelle ZV, so heißt

$$\mathbb{E}(X) := \int X \, d\mathbb{P} = \int x \, d\mathbb{P}_X(x)$$

der Erwartungswert von X.

Definition 5.5 (Varianz einer ZV)
Ist X eine ZV mit $\mathbb{E}(|X|) < \infty$ so heißt

$$\mathbb{V}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Varianz von X.

5.3 Funktionen von Zufallsvariablen

Satz 5.6 (Funktionen von Zufallsvariablen)
Sei $X : \Omega \to \mathbb{R}^n$ eine Zufallsvariable und $g : \mathbb{R}^n \to \mathbb{R}^m$ meßbar. Dann ist $g \circ X : \Omega \to \mathbb{R}^m$ ebenfalls eine Zufallsvariable.

5.4 Momente und momenterzeugende Funktionen

Definition 5.7 (Momente)
Sei X eine integrierbare Zufallsvariable und $n \in \mathbb{N}$, so daß X^n quasi-integrierbar ist. Dann heißt

$$\begin{array}{ll} m_{(n)}(X) = \mathbb{E}(|X|^n) & n\text{-tes absolutes Moment} \\ m_n(X) = \mathbb{E}(X^n) & n\text{-tes Moment und} \\ m_n^0(X) = \mathbb{E}((X - \mathbb{E}(X))^n) & n\text{-tes zentriertes Moment von } X \end{array}$$

Definition 5.8 (symmetrische Verteilung)
Sei $\mathbb{P}$ Verteilung auf $(\mathbb{R}, \mathcal{B})$.
$\mathbb{P}$ heißt symmetrisch um $a \in \mathbb{R} \quad \Leftrightarrow \quad \mathbb{P}(]-\infty, a-x]) = \mathbb{P}([a+x, \infty[)$

Satz 5.9
Sei $\mathbb{P}$ eine um a symmetrische Verteilung auf $(\mathbb{R}, \mathcal{B})$; $X \sim \mathbb{P}_X = \mathbb{P}$. Dann gilt

i) $g \in L(\mathbb{R}, \mathcal{B}, \mathbb{P}) \Rightarrow \displaystyle\int g \, d\mathbb{P}_X = \int g(x) \, d\mathbb{P}_X(x) = \int g\,(2a - x) \, d\mathbb{P}_X(x)$

ii) $m_1(X) = a$

iii) $m_{2n+1}^0(X) = 0 \quad \forall n \in \mathbb{N}$.

Definition 5.10 (momenterzeugende Funktion)
Ist X eine reelle ZV und $\mathcal{D} := \{s \in \mathbb{R} | \mathbb{E}(\exp(sX)) < \infty\}$, so heißt die Funktion

$$\begin{aligned} M : \mathcal{D} &\to \mathbb{R} \\ s &\mapsto \mathbb{E}(\exp(sX)) = \int \exp(sx) \, d\mathbb{P}_X(x) \end{aligned}$$

momenterzeugende Funktion.

Satz 5.11
Sei X eine ZV mit momenterzeugender Funktion $M : \mathcal{D} \to \mathbb{R}$. Ist $]-a, a[\subset \mathcal{D}$ für ein beliebiges $a > 0$, so gilt

i) $\mathbb{E}(X^n) < \infty$

ii) $M(s) = \sum\limits_{n=0}^{\infty} \frac{s^n}{n!} \mathbb{E}(X^n)$

iii) $\left.\dfrac{\partial^n M(s)}{\partial^n s}\right|_{s=0} = \mathbb{E}(X^n)$

Kapitel 6

Bedingte Wahrscheinlichkeit und Unabhängigkeit

6.1 Bedingte Wahrscheinlichkeit

Definition 6.1 (bedingte W'keit)
Ist $(\Omega, \mathcal{F}, \mathbb{P})$ ein W'keitsraum und $\mathbb{P}(B) > 0$ für ein $B \in \mathcal{F}$, so heißt

$$\begin{aligned} \mathbb{P}(\cdot \,|\, B) : \mathcal{F} &\to [0,1] \\ A &\mapsto \mathbb{P}(A \,|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \end{aligned}$$

die bedingte W'keitverteilung bzw. bedingte Wahrscheinlichkeit von A gegeben B.

Satz 6.2
$\mathbb{P}(\cdot \,|\, B)$ ist W'keitsmaß.

Satz 6.3
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein W'keitsraum und $A, B \in \mathcal{F}$ sowie $A_n \in \mathcal{F}, n \in \mathbb{N}$. Dann gilt

$$\begin{aligned} \text{Ist} \quad \mathbb{P}\left(\bigcap_{i=1}^{n-1} A_i\right) &> 0 \Rightarrow \\ \mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) &= \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \,|\, A_1) \cdot \ldots \cdot \mathbb{P}\left(A_n \,\middle|\, \bigcap_{i=1}^{n-1} A_i\right) \end{aligned}$$

Satz 6.4 (Satz von der totalen Wahrscheinlichkeit)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ W'keitsraum. Sei $\dot{\bigcup}_{i=1}^{\infty} A_i = \Omega$ eine disjunkte Zerlegung von Ω mit $\mathbb{P}(A_i) > 0 \quad \forall i \in \mathbb{N}$. Dann gilt

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \cdot \mathbb{P}(A \,|\, A_i) \quad \forall A \in \mathcal{F}.$$

Satz 6.5 (Satz von Bayes)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ W'keitsraum mit $\dot{\bigcup}_{i=1}^{\infty} A_i = \Omega$ einer disjunkten Zerlegung von Ω mit $\mathbb{P}(A_i) > 0 \quad \forall i \in \mathbb{N}$ und $B \in \mathcal{F}$ mit $\mathbb{P}(B) > 0$. Dann gilt

$$\mathbb{P}(A_i \mid B) = \frac{\mathbb{P}(B \mid A_i)\mathbb{P}(A_i)}{\sum\limits_{i=1}^{\infty} \mathbb{P}(A_i)\mathbb{P}(B \mid A_i)}$$

6.2 Unabhängigkeit

Definition 6.6 (Unabhängigkeit von Ereignissen)
Ereignisse $A_i \in \mathcal{F}, i \in I \neq \emptyset$, eines W'keitsraumes $(\Omega, \mathcal{F}, \mathbb{P})$ heißen stochastisch unabhängig (stu), wenn für jede endliche Teilmenge $\emptyset \neq J \subset I$ gilt

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$$

Satz 6.7

a) A, B stu, $\mathbb{P}(B) > 0 \Rightarrow \mathbb{P}(A \mid B) = \mathbb{P}(A)$

b) A, B stu $\Rightarrow A, \bar{B}$ stu.

Satz 6.8 (Borel-Cantelli-Lemma)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ W'keitsraum und $A_i, i \in \mathbb{N}$, eine Folge von Ereignissen mit $A_i \in \mathcal{F}$.
Ist $\sum\limits_{i=1}^{\infty} \mathbb{P}(A_i) < \infty$, so folgt

a) $\mathbb{P}(\limsup A_i) = 0$

Sind die A_i stu und $\sum\limits_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$, so folgt

b) $\mathbb{P}(\limsup A_i) = 1$

Kapitel 7

Mehrdimensionale Verteilungen

7.1 Mehrdimensionale Verteilungen

Definition 7.1 (n-dimensionale ZV, Verteilungsfunktion)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X : \Omega \to \mathbb{R}^n$ eine n-dimensionale reelle ZV, $X = (X_1, \ldots, X_n)$. Dann heißt die Funktion

$$\begin{array}{rcl} F_X : \mathbb{R}^n & \to & [0,1] \\ x = (x_1, \ldots, x_n) & \mapsto & \mathbb{P}(X \leq x) \\ & = & \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \\ & = & \mathbb{P}(\{\omega \in \Omega \,|\, X_i(\omega) \leq x_i \; \forall i = 1, \ldots, n\}) \end{array}$$

die n-dimensionale Verteilungsfunktion von X.

Satz 7.2
X, Y n-dimensionale ZV, $X \sim F_X, Y \sim F_Y$.

$$F_X = F_Y \Rightarrow \mathbb{P}_X = \mathbb{P}_Y.$$

Satz 7.3
Sei $X : \Omega \to \mathbb{R}^n$ eine n-dimensionale ZV und $g : \mathbb{R}^n \to \mathbb{R}^k$ $\mathcal{B}^n$-$\mathcal{B}^k$-meßbar (Baire-Funktion). Dann ist

$$g \circ X : \Omega \to \mathbb{R}^k, \quad \omega \mapsto g(X(\omega))$$

eine k-dimensionale reelle ZV mit Verteilung

$$\mathbb{P}_{g(X)}(A) = \mathbb{P}(g(X) \in A) = \mathbb{P}(\{\omega \in \Omega \,|\, g(X(\omega)) \in A\})$$

Definition 7.4 (Randverteilung)
Ist $g : \mathbb{R}^n \to \mathbb{R}$ mit $g(x) = x_j$ für ein festes $j \in \{1, \ldots, n\}$ (Projektion auf j-te Koordinate), so gilt:

$$\begin{array}{rcl} g(X) & = & X_j \\ \mathbb{P}_{X_j}(A) & = & \mathbb{P}_X(\{x \in \mathbb{R}^n : x_j \in A\}) = \mathbb{P}(X_j \in A) \text{ für } A \in \mathcal{B} \end{array}$$

$\mathbb{P}_{X_j}$ heißt Randverteilung oder marginale Verteilung von X_j.

7.2 Transformationssatz für Dichten

Satz 7.5
Sei $X : \Omega \to \mathbb{R}$ ZV mit stetiger Verteilungsfunktion F_X und Dichte $f_X(x) = \dfrac{\partial F_X(x)}{\partial x}$ bezüglich

des Lebesgue-Maßes λ. Weiterhin sei $g : \mathbb{R} \to \mathbb{R}$ bijektiv und stetig differenzierbar, $\frac{\partial g(x)}{\partial x} \neq 0$ mit $h = g^{-1}$.

$\Rightarrow g \circ X$ hat Dichte $\quad f_{g\circ X}(y) = f_X(h(y)) \cdot \left|\frac{\partial h(y)}{\partial y}\right| \quad$ bezüglich λ.

Satz 7.6 (Trafo für injektive g)
Sei $X : \Omega \to \mathbb{R}$ mit Verteilungsfunktion F_X und stetiger Dichte f_X sowie $g : \mathbb{R} \to \mathbb{R}$ eine Baire-Funktion. Sei $\mathbb{R} \supset I = \dot{\bigcup\limits_{m\in M}} I_m$ eine disjunktive Vereinigung offener Intervalle ($I_m =]a_m, b_m[$) derart, daß gilt:

a) F_X ist stetig differenzierbar auf I:

$$f_X(x) = \frac{\partial F_X(x)}{\partial x}$$

b) $\mathbb{P}(X \in I) = 1$

c) Sei $g_m = g|_{I_m}$; $\quad g_m : I_m \to \mathbb{R}$ die Einschränkung von g auf I_m mit

   1) g_m ist stetig differenzierbar und bijektiv
   2) $\frac{\partial g_m(x)}{\partial x} \neq 0$

   Dann gilt: $g \circ X$ hat die Dichte

$$f_{g\circ X}(y) = \sum_{m\in M} v_m(y)$$

   wobei $v_m(y) = \begin{cases} f_X(h_m(y)) \cdot \left|\frac{\partial h_m(y)}{\partial y}\right| & y \in g(I_m) \\ 0 & \text{sonst} \end{cases}$
   und $h_m = g_m^{-1}$

Satz 7.7 (Verallgemeinerungen auf den n-dimensionalen Fall)
Sei $X : \Omega \to \mathbb{R}^n$ ZV mit Dichte f_X bzgl. λ^n.
Sei $g : \mathbb{R}^n \to \mathbb{R}^n$ Baire-Funktion.
Ferner sei $G_m \in \mathcal{B}^n, m \in M$ derart, daß

i) $\mathbb{P}\left(X \in \dot{\bigcup\limits_{m\in M}} G_m\right) = 1$

ii) $g_m := g|_{G_m}$ bijektiv und stetig differenzierbar.

Sei $H_m(y) = \underbrace{\left(\frac{\partial h_m(y)}{\partial y_1 \partial y_2 \cdots \partial y_n}\right)}_{=:J \text{ Jacobi-Matrix von } h_m}$ und $h_m = g_m^{-1}$.

Dann gilt

$$f_{g\circ X}(y) = \sum_{m\in M} v_m(y)$$

mit $v_m(y) = \begin{cases} f_X(h_m(y)) \cdot |\det(H_m(y))| & \text{falls } h_m(y) \in G_m \\ 0 & \text{sonst} \end{cases}$

7.3 Unabhängigkeit von Zufallsvariablen

Definition 7.8

i) Eine Familie von Mengensystemen $\mathcal{F}_i \subset \mathcal{F},\ i \in I \neq \emptyset$, heißt stochastisch unabhängig, wenn dies für jede Familie von Ereignissen $A_i \in \mathcal{F}_i,\ i \in I$, für alle $i \in I$ gilt. Vgl. Definition 6.6

ii) Eine Familie von Zufallsvariablen $X_i : \Omega \to \Omega_i$, auf einem W'keitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ heißt stochastisch unabhängig, wenn dies für die Familie der von den Zufallsvariablen X_i erzeugten σ-Algebra
$$\sigma(X_i) = \sigma(X_i^{-1}(\mathcal{F}_i)) \quad i \in I$$
gilt.

Satz 7.9
Ist $X = (X_1, \ldots, X_n)$ reelle n-dimensionale ZV, so sind $X_1, \ldots, X_n$ stu $\quad\Leftrightarrow\quad \forall c \in \mathbb{R}^n$ gilt

$$\begin{aligned} \mathbb{P}(X \le c) &= \mathbb{P}(X_1 \le c_1, \ldots, X_n \le c_n) \\ &= \prod_{i=1}^n \mathbb{P}(X_i \le c_i) \\ \text{bzw.}\quad F_X(c) &= \prod_{i=1}^n F_{X_i}(c_i) \end{aligned}$$

Ist X diskret, so gilt: $X_1, \ldots, X_n$ stu $\quad\Leftrightarrow\quad \forall x \in T_1 \times T_2 \times \ldots, \times T_n \quad (T_i$ Träger von $X_i)$ gilt

$$\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^n \mathbb{P}(X_i = x_i)$$

Satz 7.10
Sind $X_1, \ldots, X_n$ reelle unabhängige ZV mit Dichten $f_{X_1}, \ldots, f_{X_n}$ bzgl. λ, dann und genau dann ist die gemeinsame Dichte von $X = (X_1, \ldots, X_n)$ das Produkt der einzelnen Dichten, d.h.

$$f_X(x) = \prod_{i=1}^n f_{X_i}(x_i)$$

ist Dichte bzgl. λ^n.

Satz 7.11
Seien $X_1 : \Omega \to \mathbb{R}^{n_1}$ und $X_2 : \Omega \to \mathbb{R}^{n_2}$ zwei n_i-dimensionale stochastisch unabhängige Zufallsvariablen $(i = 1, 2)$ und $h_i : \mathbb{R}^{n_i} \to \mathbb{R}^{m_i}$ Baire-Funktionen.

$$\Rightarrow \quad Y_1 = h_1 \circ X_1 \text{ und } Y_2 = h_2 \circ X_2 \text{ stu.}$$

Satz 7.12 (Erwartungswert des Produkts unabhängiger ZV)
Sind $X_1, \ldots, X_n$ unabhängige integrierbare ZV, so folgt

$$\mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i)$$

(Die Umkehrung folgt nicht!)

Satz 7.13
Seien $X_1, \dots, X_n$ unabhängige reelle ZV, deren momenterzeugende Funktionen $M_1, \dots, M_n$ alle auf dem Intervall $]-a, a[$ definiert sind.
Sei $S_n := \sum_{i=1}^{n} X_i$, so ist auch

$$M(s) = \mathbb{E}(\exp(s\, S_n))$$

auf $]-a, a[$ definiert und es gilt

$$M(s) = \prod_{i=1}^{n} M_i(s), \quad s \in]-a, a[$$

7.4 Einige Ungleichungen

Satz 7.14 (Markow- und Tschebyschow-Ungleichungen)
Sei $X : \Omega \to \mathbb{R}$ eine reelle ZV. Dann gilt für jedes $\epsilon > 0$:

$$\begin{aligned} \mathbb{P}(|X| \geq \epsilon) &\leq \frac{1}{\epsilon^n} \int |X^n|\, I_{\{|X| \geq \epsilon\}}\, d\mathbb{P} \\ &\leq \frac{1}{\epsilon^n} \mathbb{E}(|X|^n) \end{aligned}$$

Insbesondere $(n = 1)$

$$\mathbb{P}(|X| \geq \epsilon) \leq \frac{1}{\epsilon} \mathbb{E}(|X|) \quad \text{(Markow-Ungleichung)}$$

und für $\mathbb{E}(|X|) < \infty$

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon) \leq \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\epsilon^2} = \frac{\mathbb{V}(X)}{\epsilon^2} \quad \text{(Tschebyschow-Ungleichung).}$$

Satz 7.15 (Jensen'sche Ungleichung)
Sei $f : I \to \mathbb{R}$ eine konvexe Funktion auf einem Intervall $I \subset \mathbb{R}$ und $X : \Omega \to I$ eine integrierbare ZV. Dann gilt

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$$

7.5 Der mehrdimensionale Erwartungswert und die Kovarianzmatrix

Definition 7.16
Sei $X : \Omega \to \mathbb{R}^n$, $X = (X_1, \dots, X_n)$ eine n-dimensionale ZV, dann heißt

$$\mathbb{E}(X) := (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))$$

der (n-dimensionale) Erwartungswert von X und

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))(X - \mathbb{E}(X))^\top)$$

die Kovarianzmatrix von X.

Satz 7.17
Sei $X : \Omega \to \mathbb{R}^n$ n-dimensionale ZV, $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Dann gilt

i) $\mathbb{E}(AX + b) = A\,\mathbb{E}(X) + b$

ii) $\mathbb{V}(AX+b) = A\,\mathbb{V}(X)\,A^\top$

iii) $\mathbb{V}(X)$ psd

Satz 7.18
X, Y ZV, $X_1, \ldots, X_n$ ZV

i) $\operatorname{Cov}(X,Y) = \mathbb{E}(X\,Y) - \mathbb{E}(X)\,\mathbb{E}(Y)$

ii) X, Y stu $\quad\Rightarrow\quad \operatorname{Cov}(X,Y) = 0$

iii)

$$\begin{aligned}\mathbb{V}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \mathbb{V}(X_i) + 2\sum_{1\le i<j\le n} \operatorname{Cov}(X_i, X_j)\\ &\overset{X_1,\ldots,X_n \text{ stu}}{=} \sum_{i=1}^n \mathbb{V}(X_i)\end{aligned}$$

Definition 7.19 (Korrelationskoeffizient)
Seien X, Y ZV mit $\mathbb{V}(X) < \infty$, $\mathbb{V}(Y) < \infty$. Dann heißt

$$\rho(X,Y) := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)\,\mathbb{V}(Y)}}$$

der Korrelationskoeffizient von X und Y.

Satz 7.20

$$|\rho(X,Y)| \le 1$$

Satz 7.21 (Standardisierung)
Sei $X = (X_1, \ldots, X_n)$ n-dimensionale ZV mit $\mathbb{E}(X) = \mu$ und $\mathbb{V}(X) = \Sigma$. Dann existiert $B \in \mathbb{R}^{n,n}$ mit $Y = B^{-1}(X-\mu)$ und $\mathbb{E}(Y) = 0_n$ sowie $\mathbb{V}(Y) = \mathbf{E}_n$.

Kapitel 8

Einige spezielle Verteilungen und deren Eigenschaften

8.1 Diskrete Verteilungen

Satz 8.1 (Erwartungswert und Varianz der $B(n,p)$)
Seien $X_1, \dots, X_n$ stu mit $X_i \sim B(1,p),\ i = 1, \dots, n$. Dann gilt für $X = \sum_{i=1}^{n} X_i$

$$\begin{aligned} \mathbb{E}(X) &= n \cdot p \\ \mathbb{V}(X) &= n \cdot p \cdot (1-p) \end{aligned}$$

Satz 8.2 (Zusammenhang zwischen $B(1,p)$ und $B(n,p)$)
Seien $X_1, \dots, X_n$ stu mit $X_i \sim B(1,p)$. Dann gilt:

$$X = \sum_{i=1}^{n} X_i \sim B(n,p).$$

Korollar 8.3
$X_1 \sim B(n_1,p)$, $X_2 \sim B(n_2,p)$ stu. $\Rightarrow X_1 + X_2 \sim B(n_1 + n_2, p)$ Summenstabilität

Definition 8.4 (Hypergeometrische Verteilung)
Modell: Ziehen ohne Zurücklegen
$X \sim H(n,K,N)$,

$$\mathbb{P}(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Satz 8.5 (Erwartungswert und Varianz von $X \sim H(n,K,N)$)
$X \sim H(n,K,N)$

$$\begin{aligned} \mathbb{E}(X) &= n \cdot \frac{K}{N} \\ \mathbb{V}(X) &= n \cdot \frac{K}{N}\left(1 - \frac{K}{N}\right) \cdot \left(\frac{N-n}{N-1}\right) \end{aligned}$$

Satz 8.6 (Zusammenhang zwischen $B(n,p)$ und $H(n,K,N)$)
Seien $K_m, N_m \in \mathbb{N}$ mit $0 \leq K_m \leq N_m$ Folgen $(m \in \mathbb{N})$ mit $K_m \to \infty, N_m \to \infty$ und $p \in [0,1], n \in \mathbb{N}$ fest. Dann gilt mit $\frac{K_m}{N_m} \to p$ für $m \to \infty$

$$\lim_{m\to\infty} \mathbb{P}(H_m = k) = \mathbb{P}(B = k) \quad \forall k \in \mathbb{N}$$

für $H_m \sim H(n, K_m, N_m)$ und $B \sim B(n, p)$.

Satz 8.7 (Zusammenhang zwischen $B(n,p)$ und $P(\lambda)$)
Sei $p_n \in [0,1], n \in \mathbb{N}$, eine Folge und $\lambda_n = n \cdot p_n \to \lambda > 0$ für $n \to \infty$. Dann gilt

$$b(k, n, p_n) \overset{n\to\infty}{\to} e^{-\lambda} \frac{\lambda^k}{k!} \quad \forall k \in \mathbb{N}.$$

Definition 8.8 (Geometrische Verteilung)
Die diskrete Verteilung mit Parameter $p \in [0,1]$ und Dichte

$$\mathbb{P}(X = x) = (1-p)^{x-1}\, p \quad x \in \mathbb{N}$$

heißt geometrische Verteilung, kurz $X \sim G(p)$, und beschreibt die Wahrscheinlichkeit der Anzahl von Versuchen bis zum ersten Erfolg.

Satz 8.9 (Eigenschaften der Geometrischen Verteilung)
$X \sim G(p)$

(a) $\mathbb{E}(X) = \frac{1}{p}$ und $\mathbb{V}(X) = \frac{1-p}{p^2}$

(b) Gedächtnislosigkeit $\mathbb{P}(X = x + x_0 | X > x_0) = \mathbb{P}(X = x)$

Definition 8.10 (Negative Binomialverteilung)
Die diskrete Verteilung mit Parametern $p \in [0,1]$ und $n \in \mathbb{N}$ sowie der Dichte

$$\mathbb{P}(X = x) = \binom{x-1}{n-1} p^n \,(1-p)^{x-n} \quad \text{für } x \in \mathbb{N}, x \geq n$$

heißt negative Binomialverteilung, kurz $X \sim B^{-1}(n,p)$.

Satz 8.11 (Zusammenhang zwischen $G(p)$ und $B^{-1}(n,p)$)
Seien $X_1, \ldots, X_n$ stu mit $X_i \sim G(p), i = 1, \ldots, n$. Dann gilt

$$X = \sum_{i=1}^{n} X_i \sim B^{-1}(n,p)$$

Korollar 8.12
$X \sim B^{-1}(n,p) \Rightarrow \mathbb{E}(X) = \frac{n}{p}$
$X = \sum\limits_{i=1}^{n} X_i \; X_i \sim G(p)$ stu.
$\mathbb{E}(X) = n \cdot \mathbb{E}(X_i) = \frac{n}{p}$ und $\mathbb{V}(X) = \frac{n(1-p)}{p^2}$
Negative Binomialverteilung ist nicht gedächtnislos.

Definition 8.13 (Multinomialverteilung)
Sei $n \in \mathbb{N}$ und $p_1, \ldots, p_k \in [0,1]$ mit $\sum\limits_{j=1}^{k} p_j = 1$. Sei $Y = (Y_1, \ldots, Y_k)$ eine k-dimensionale diskrete ZV mit Dichte

$$\begin{aligned} f_Y(y) &= \mathbb{P}(Y_j = y_j \quad \forall j = 1, \ldots, k) \\ &= \begin{cases} \dfrac{n!}{y_1! \cdots y_k!} p_1^{y_1} \cdots p_k^{y_k} & \text{falls } \sum_{j=1}^{k} y_j = n \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Dann heißt die zugehörige Verteilung Multinomialverteilung mit Parametern n und $p_1, \ldots, p_k$: $Y \sim M(n, p_1, \ldots, p_k)$.

Satz 8.14 (Eigenschaften der Multinomialverteilung)

i) $Y_i \sim B(n, p_i)$

ii) $\mathbb{E}(Y) = n\,(p_1, \ldots, p_k)$

iii) $\mathbb{V}(Y) = n \begin{pmatrix} p_1\,(1-p_1) & -p_1\,p_2 & \cdots & -p_1\,p_k \\ -p_1\,p_2 & p_2\,(1-p_2) & \cdots & -p_2\,p_k \\ \vdots & & & \\ -p_1\,p_k & -p_2\,p_k & \cdots & p_k\,(1-p_k) \end{pmatrix}$

8.2 Die univariate Normalverteilung

Bekannt: $X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \mathbb{E}(X) = \mu, \quad \mathbb{V}(X) = \sigma^2$

Satz 8.15 (Lineare Transformation)

$$X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad Y = aX + b \sim N(a\mu + b, a^2\sigma^2) \quad \text{für } a, b \in \mathbb{R}, a \neq 0.$$

Korollar 8.16

$$X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad Y = \frac{X - \mu}{\sigma} \sim N(0, 1).$$

Satz 8.17
$X \sim N(\mu, \sigma^2) \Rightarrow$
$m_r^0(X) = \begin{cases} 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (r-1)\,\sigma^r & r = 2k, k \in \mathbb{N} \\ 0 & \text{sonst} \end{cases}$

Satz 8.18 (Additivität der Normalverteilung)
Seien $X_1 \sim N(\mu_1, \sigma_1^2)$ und $X_2 \sim N(\mu_2, \sigma_2^2)$ stu

$$\Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Korollar 8.19
$X_1, \ldots, X_n$ stu, $X_i \sim N(\mu_i, \sigma_i^2)$ für $i = 1, \ldots, n$. Dann gilt

i) $\sum X_i \sim N(\sum \mu_i, \sum \sigma_i^2)$

ii) $\mu_1 = \ldots = \mu_n, \sigma_1^2 = \ldots = \sigma_n^2 \quad \Rightarrow \quad \bar{X} = \frac{1}{n} \sum X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Korollar 8.20
$X_1 \sim N(\mu_1, \sigma_1^2), X_2 \sim N(\mu_2, \sigma_2^2)$ stu. $\Rightarrow \quad aX_1 + bX_2 + c \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2)$

8.3 Die k-dimensionale Normalverteilung

Definition 8.21 (k-dimensionale Normalverteilung)
Sei $X : \Omega \to \mathbb{R}^k$. $X = (X_1, \ldots, X_k)$ heißt k-dimensional (multivariat) standardnormalverteilt, wenn X eine stetige Verteilung mit Dichte

$$f_X(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^k \exp\left(-\frac{1}{2}\sum x_i^2\right) \quad x = (x_1, \ldots, x_k) \in \mathbb{R}^k$$

besitzt: $X \sim N_k(0, I_k)$.

Satz 8.22

$$X \sim N_k(0, I_k) \quad \Leftrightarrow \quad X_1, \ldots, X_k \text{ stu}, X_i \sim N(0,1)\ \forall i = 1, \ldots, k.$$

Definition 8.23
Sei $X \sim N_k(0, I_k)$ und $A \in \mathbb{R}^{p,k}$ sowie $\mu \in \mathbb{R}^p$. Dann heißt

$$Y = A\,X + \mu$$

p-dimensional normalverteilt:

$$Y \sim N_p(\mu, \Sigma) \quad \text{mit } \Sigma = A\,A^\top$$

Satz 8.24

$$Y \sim N_p(\mu, \Sigma) \quad \Rightarrow \quad \begin{aligned} \mathbb{E}(Y) &= \mu \\ \mathbb{V}(Y) &= \Sigma \end{aligned}$$

Satz 8.25
Wenn $A \in \mathbb{R}^{k,k}$ invertierbar, dann hat $Y \sim N_k(\mu, \Sigma)$ die Dichte $f_Y : \mathbb{R}^k \to \mathbb{R}$

$$f_Y(y) = \frac{1}{\sqrt{(2\pi)^k \det(\Sigma)}} \exp\left(-\frac{(y-\mu)^\top\,\Sigma^{-1}\,(y-\mu)}{2}\right)$$

mit $\Sigma = AA^\top$.

Satz 8.26

a) $Y \sim N_k(\mu, \Sigma) \quad \Rightarrow \quad Y_i \sim N(\mu_i, \Sigma_{ii}) \quad \forall i = 1, \ldots, k$

b) $Y \sim N_k(\mu, \Sigma) \quad \Leftrightarrow \quad n^\top Y \sim N_1(n^\top \mu, n^\top \Sigma n) \quad \forall n \in \mathbb{R}^k$

c) $(Y_1, Y_2) \sim N_2(\mu, \Sigma)$ mit $\mu = (\mu_1, \mu_2)$ und $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

$$Y_1, Y_2 \text{ stu} \quad \Leftrightarrow \quad \Sigma_{12} = \Sigma_{21} = 0$$

8.4 Gamma- und χ^2-Verteilung

Definition 8.27 (Γ-Funktion)

$$\begin{aligned} \Gamma : \mathbb{R}^+ &\to \mathbb{R}^+ \\ \Gamma(z) &= \int_0^\infty t^{z-1}\, e^{-t}\, dt \end{aligned}$$

Satz 8.28 (Eigenschaften der Γ-Funktion)

- $\Gamma(z+1) = z\,\Gamma(z)$
- $\Gamma(1) = 1$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

- $\Gamma(m+1) = m! \quad m \in \mathbb{N}$
- $\Gamma\left(\frac{k+1}{2}\right) = \frac{k-1}{2} \cdot \frac{k-3}{2} \cdot \ldots \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi}$ falls k gerade.

Definition 8.29
Eine ZV X mit stetiger Dichte

$$f_X(x; k, \lambda) = \frac{\lambda}{\Gamma(k)} (\lambda x)^{k-1} \exp(-\lambda x) \cdot I_{(0,\infty)}(x)$$

Mit $k > 0$ und $\lambda > 0$ heißt Γ-verteilt: $X \sim \Gamma(k, \lambda)$.

Satz 8.30

$$\Gamma(k = 1, \lambda) = Exp(\lambda)$$

Satz 8.31
$\mathbb{E}(X) = \dfrac{k}{\lambda}$, $\mathbb{V}(X) = \dfrac{k}{\lambda^2}$ und $M_X(s) = \left(\dfrac{\lambda}{\lambda - s}\right)^k$ für $s < \lambda$.

Definition 8.32
Eine ZV X mit stetiger Dichte

$$f_X(x; k) = \frac{1}{\Gamma\left(\frac{k}{2}\right)} \left(\frac{1}{2}\right)^{\frac{k}{2}} x^{\frac{k}{2}-1} \exp\left(-\frac{1}{2}x\right) \cdot I_{(0,\infty)}(x)$$

heißt χ^2-verteilt mit k Freiheitsgraden: $X \sim \chi^2_k$.

Satz 8.33
$\mathbb{E}(X) = k$, $\mathbb{V}(X) = 2k$, $M_X(s) = \left(\dfrac{1}{1-2s}\right)^{\frac{k}{2}}$ $s < \frac{1}{2}$

Satz 8.34
$X_1, \ldots, X_n$ stu mit $X_i \sim N(0,1)$. Dann

$$Y = \sum_{i=1}^{n} X_i^2 \sim \chi^2_n.$$

8.5 t- und F-Verteilung

Satz 8.35
Sei $U \sim \chi^2_m$ und $V \sim \chi^2_n$ stu. Dann hat die Zufallsvariable

$$X = \frac{U/m}{V/n}$$

eine stetige Verteilung mit Dichte

$$f_X(x; m, n) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{x^{\frac{m-2}{2}}}{\left(1 + \left(\frac{m}{n}\right) \cdot x\right)^{\frac{m+n}{2}}} \cdot I_{(0,\infty)}(x)$$

Definition 8.36 (F-Verteilung)
Eine ZV X mit Dichte f_X heißt F-verteilt mit Freiheitsgraden m und n: $X \sim F_{m,n}$.

Satz 8.37

$$\begin{aligned} X \sim F_{m,n} \Rightarrow \mathbb{E}(X) &= \frac{n}{n-2} \text{ für } n > 2 \text{ und} \\ \mathbb{V}(X) &= \frac{2\,n^2\,(m+n-2)}{m\,(n-2)^2\,(n-4)} \text{ für } n > 4 \end{aligned}$$

Satz 8.38
Sei $Z \sim N(0,1)$ und $U \sim \chi_k^2$ stu. Dann hat

$$X = \frac{Z}{\sqrt{U/k}}$$

eine Verteilung mit Dichte

$$f_X(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \frac{1}{\sqrt{k\pi}} \frac{1}{\left(1+\frac{x^2}{k}\right)^{\frac{k+1}{2}}}$$

Definition 8.39 (t-Verteilung)
Eine ZV X mit Dichte f_X heißt t-verteilt mit k Freiheitsgraden: $X \sim t_k$.

Satz 8.40

a) $\mathbb{E}(X) = 0$

b) $\mathbb{V}(X) = \dfrac{k}{k-2}$

Kapitel 9

Konvergenz und Wahrscheinlichkeit

9.1 Konvergenzarten

Definition 9.1 (Konvergenzarten)
Sei $(\Omega, \mathcal{F}, \mathbb{P})$ W'keitsraum und $X_i : \Omega \to \mathbb{R}$ ZV, $i = 1, 2, \ldots$. Dann konvergiert die Folge $(X_n)_{n\in\mathbb{N}}$ gegen eine ZV $X : \Omega \to \mathbb{R}$

a) <u>fast sicher</u>, wenn

$$\mathbb{P}(\{\omega \in \Omega \,|\, X_n(\omega) \to X(\omega) \text{ für } n \to \infty\}) = 1;$$

$$X_n \overset{\text{f.s.}}{\to} X,\ X_n \overset{\text{a.s.}}{\to} X,\ X_n \overset{\text{a.e.}}{\to} X,\ X_n \to X \text{ wp } 1.$$

b) <u>im r-ten Moment</u>, $r \geq 1$, wenn

$$\mathbb{E}(|X_n|^r) < \infty\ \forall\, n \quad \text{und} \quad \mathbb{E}(|X_n - X|^r) \to 0 \quad \text{für } n \to \infty;$$

$$X_n \overset{r}{\to} X$$

c) <u>in Wahrscheinlichkeit</u>, falls

$$\mathbb{P}(|X_n - X| > \epsilon) \to 0 \quad \text{für } n \to \infty\ \forall \epsilon > 0;$$

$$X_n \overset{\mathbb{P}}{\to} X$$

d) <u>in Verteilung</u>, falls

$$\lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$$

und alle Stetigkeitsstellen von $F_X(x) = \mathbb{P}(X \leq x)$;

$$X_n \overset{D}{\to} X$$

Satz 9.2
Sei $(X_n)_{n\in\mathbb{N}}$, Folge von ZV $X_i : \Omega \to \mathbb{R}$ und $X : \Omega \to \mathbb{R}$ ZV. Dann gilt

a) $X_n \overset{\mathbb{P}}{\to} X \quad \Rightarrow \quad X_n \overset{D}{\to} X$

b) $X_n \overset{\text{f.s.}}{\to} X \quad \Rightarrow \quad X_n \overset{\mathbb{P}}{\to} X$

c) $X_n \overset{r}{\to} X \quad \Rightarrow \quad X_n \overset{\mathbb{P}}{\to} X \quad$ für $r \geq 1$

d) $X_n \stackrel{r}{\to} X \quad \Rightarrow \quad X_n \stackrel{s}{\to} X \quad r > s \geq 1$

Satz 9.3

a) $X_n \stackrel{D}{\to} c$ konstant $\quad \Rightarrow \quad X_n \stackrel{\mathbb{P}}{\to} c$

b) $X_n \stackrel{\mathbb{P}}{\to} X$ und $\mathbb{P}(|X_n| \leq k) = 1 \quad \forall n$ und ein $k \in \mathbb{R}^+$

$\Rightarrow X_n \stackrel{r}{\to} X \quad \forall r \geq 1$

c) $P_n(\epsilon) = \mathbb{P}(|X_n - X| > \epsilon)$ mit $\sum_n P_n(\epsilon) < \infty \quad \forall \epsilon > 0$

$\Rightarrow X_n \stackrel{\text{f.s.}}{\to} X$

9.2 Konvergenz in Wahrscheinlichkeit

Satz 9.4 (Schwaches Gesetz der großen Zahlen)
Seien $X_1, \ldots, X_n$ unabhängig und identisch verteilt mit $\mathbb{E}(X_i) = \mu$ und $\mathbb{V}(X_i) = \sigma^2 < \infty \quad \forall i = 1, \ldots, n$. Dann gilt

$$\bar{X} := \frac{1}{n} \sum_{i=1}^{n} X_i \stackrel{\mathbb{P}}{\to} \mu.$$

Satz 9.5 (Eindeutigkeit der Konvergenz)

$$X_n \stackrel{\mathbb{P}}{\to} X \quad \text{und} \quad X_n \stackrel{\mathbb{P}}{\to} \widetilde{X} \quad \Rightarrow \quad \mathbb{P}(X = \widetilde{X}) = 1$$

Satz 9.6
Seien A_n und B_n Folgen von ZV mit $A_n \stackrel{\mathbb{P}}{\to} a$ und $B_n \stackrel{\mathbb{P}}{\to} b$. Dann gilt

a) $A_n + B_n \stackrel{\mathbb{P}}{\to} a + b$
$A_n - B_n \stackrel{\mathbb{P}}{\to} a - b$

b) $A_n \cdot B_n \stackrel{\mathbb{P}}{\to} a \cdot b$

c) $\frac{A_n}{B_n} \stackrel{\mathbb{P}}{\to} \frac{a}{b} \quad \forall b \neq 0$

Satz 9.7
Sei X_n eine Folge von ZV mit $X_n \stackrel{\mathbb{P}}{\to} c$ und $f : \mathbb{R} \to \mathbb{R}$ eine in $c \in \mathbb{R}$ stetige Funktion

$$\Rightarrow f(X_n) \stackrel{\mathbb{P}}{\to} f(c).$$

9.3 Konvergenz in Verteilung

Satz 9.8
Sei $(X_n)_{n \in \mathbb{N}}$ eine Folge von ZV mit Verteilungsfunktion $X_n \sim F_n$. Dann existiert $X \sim F$ mit $X_n \stackrel{D}{\to} X$ genau dann, wenn $\forall \epsilon > 0$ eine Konstante $k \in \mathbb{R}^+$ und $n_0 \in \mathbb{N}$ existiert, so daß

$$\mathbb{P}(|X_n| \leq k) \geq 1 - \epsilon \quad \forall n \geq n_0.$$

Satz 9.9

$$X_n \stackrel{D}{\to} X,\ a, b \neq 0 \in \mathbb{R} \text{ konstant} \quad \Rightarrow \quad b\,X_n + a \stackrel{D}{\to} b\,X + a.$$

Satz 9.10
$X_n \xrightarrow{D} X \quad \Leftrightarrow \quad \mathbb{E}(g(X_n)) \to \mathbb{E}(g(X))$

$\forall$ beschränkten und stetigen Funktionen $g : \mathbb{R} \to \mathbb{R}$.

Satz 9.11 (Slutsky's Theorem)

$$X_n \xrightarrow{D} X,\ A_n \xrightarrow{\mathbb{P}} a \text{ und } B_n \xrightarrow{\mathbb{P}} b \quad \Rightarrow \quad A_n + B_n\, X_n \xrightarrow{D} a + b\, X$$

Lemma 9.12

$$X_n \xrightarrow{D} X \quad \Leftrightarrow \quad M_{X_n}(s) \to M_X(s) \quad s \in \mathcal{D} \quad \text{vgl Satz 5.10}$$

Satz 9.13 (Zentraler Grenzwertsatz)
Seien $X_i, i = 1, \ldots$ u.i.v. mit $\mathbb{E}(X_i) = \mu$ und $\mathbb{V}(X_i) = \sigma^2 < \infty\ \forall i = 1, \ldots$. Dann gilt

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{D} N(0, 1)$$

bzw.

$$\sqrt{n}\,(\bar{X} - \mu) \xrightarrow{D} N(0, \sigma^2)$$

Satz 9.14 (Berry-Esseen)
Es seien $X_1, \ldots, X_n$ u.i.v., $X_i \sim F$ mit $\mathbb{E}(X_i) = \mu,\ \mathbb{V}(X_i) = \sigma^2,\ \mathbb{E}(X_i^3) < \infty \quad \Rightarrow$

$\exists\, c$ (unabhängig von F) mit

$$|G_n(x) - \Phi(x)| \le \frac{c}{\sqrt{n}} \frac{\mathbb{E}\left(|X_1 - \mu|^3\right)}{\sigma^3}$$

wobei $\quad G_n(x) = \mathbb{P}\left(\sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right) \le x\right)$.

Satz 9.15
Es seien $X_1, \ldots, X_n$ u.i.v. mit $X_i \sim F$, $\mathbb{E}(X_i) = \mu,\ \mathbb{V}(X_i) = \sigma^2,\ \mathbb{E}(X_i^3) < \infty$ und k-tem zentralen Moment $\mu_k^0 = \mathbb{E}((X_i - \mu)^k)$. Dann gilt:

$$\begin{aligned} G_n(x) &= \mathbb{P}\left(\sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right) \le x\right) \\ &= \Phi(x) + \underbrace{\frac{\mu_3^0}{6\,\sigma^3\sqrt{n}}\,(1 - x^2)\,\varphi(x)}_{\text{erste Edgeworth-Korrektur}} + o\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

Satz 9.16 (Δ-Methode)
Falls $\quad \sqrt{n}\,(X_n - \nu) \xrightarrow{D} N(0, \tau^2)$

dann auch $\quad \sqrt{n}\,(f(X_n) - f(\nu)) \xrightarrow{D} N\left(0, \tau^2 \cdot \frac{\partial f(\nu)^2}{\partial \nu}\right)$

falls $f : \mathbb{R} \to \mathbb{R}$ in ν differenzierbar ist und $\frac{\partial f(\nu)}{\partial \nu} \ne 0$.

Definition 9.17
Seien $X_1, \ldots, X_n$ u.i.v. mit Verteilungsfunktion F. Dann heißt

$$\widehat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(X_i)$$

die empirische Verteilungsfunktion von $X_1, \dots, X_n$.

Satz 9.18 (Satz von Glivenko-Cantelli)

a) $\widehat{F}(x) \xrightarrow{\text{f.s.}} F(x) \quad \forall x \in \mathbb{R}$

b) $\sqrt{n}\,(\widehat{F}_n(x) - F(x)) \xrightarrow{D} N(0, F(x)(1-F(x))) \quad \forall x \in \mathbb{R}$

c) $\sup\limits_{x\in\mathbb{R}} |\widehat{F}_n(x) - F(x)| \xrightarrow{\text{f.s.}} 0$